

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 16:10:44 ; Search time 32.86 Seconds

(without alignments)
113.650 Million cell updates/sec

Title: US-09-331-631A-25_COPY_31_85

Sequence: 1 ENPKHNKCLQSCNSERSYR.....EECEGEIPRRPRQHPER 55

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.66:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	315	100.0	605	1 FMSYBA	beta-conglycinin a
2	315	100.0	605	2 S20007	beta-conglycinin a
3	259.5	82.4	639	2 B24810	beta-conglycinin a
4	65.5	20.8	774	2 T39539	alpha-amylase homo
5	63.5	20.2	373	2 T21955	hypothetical prote
6	63.5	20.2	414	2 T21954	hypothetical prote
7	63	20.0	663	2 T26835	hypothetical prote
8	62.5	19.8	588	1 FWCNAB	alpha-globulin B p
9	62.5	19.8	3051	2 S42373	hypothetical prote
10	62	19.7	389	2 T49640	transcription fact
11	62	19.7	390	2 S44285	EAR2 protein - mou
12	62	19.7	462	2 T00708	violaxanthin de-ep
13	61	19.4	147	2 T26225	hypothetical prote
14	61	19.4	1952	2 T48814	hypothetical prote
15	60	19.0	1646	2 JH0422	voltage-dependent
16	60	19.0	2203	2 T42742	voltage-dependent
17	59.5	18.9	174	2 B0IC1T	Balbiant ring-1 ch
18	59	18.7	199	2 T28981	hypothetical prote
19	59	18.7	281	2 D29960	Balbiant ring 6 ch
20	59	18.7	344	2 S34153	msl101-1 protein -
21	59	18.7	713	2 A35502	major surface-labe
22	59	18.7	1390	2 S51364	sperm tail-specific
23	58.5	18.6	411	2 T51285	hepatocyte growth
24	58.5	18.6	1610	2 A46227	voltage-dependent
25	58	18.4	81	2 A45320	transglutaminase s
26	58	18.4	478	2 T03750	violaxanthin de-ep
27	58	18.4	798	2 S01659	integrin beta-1 ch
28	58	18.4	799	1 TJMSEB	fibronectin recept
29	58	18.4	799	2 JCA126	Integrin Beta O119

30	58	18.4	803	1 IUCH3	Integrin, band 3 p
31	58	18.4	2161	2 JH0564	calcium channel al
32	58	18.4	2181	2 A38198	calcium channel al
33	57.5	18.3	131	1 BGB02	spermatid transiti
34	57.5	18.3	179	2 S28256	NADH dehydrogenase
35	57.5	18.3	231	2 T14801	MADS box protein M
36	57.5	18.3	337	2 T27635	homeobox protein c
37	57.5	18.3	422	2 T51199	hypothetical prote
38	57.5	18.3	542	1 A54963	transcription fact
39	57.5	18.3	566	2 S22477	vicilin precursor
40	57.5	18.3	1300	2 A36502	insulin receptor-r
41	57	18.1	223	2 S43535	zev-2 protein - zeb
42	57	18.1	403	2 S02709	ear-2 protein - hu
43	57	18.1	511	2 S24345	Balbiant ring 1 pr
44	57	18.1	572	2 T29880	hypothetical prote
45	57	18.1	633	2 T47524	hypothetical prote

ALIGNMENTS

RESULT 1

FMSYBA
beta-conglycinin alpha chain precursor - soybean

C:Species: Glycine max (soybean)

C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jul-1999

C:Accession: S14681; S74124; S06714

R:Sebastiani, F.L.; Farrell, L.B.; Schuller, M.A.; Beachy, R.N.

Plant Mol. Biol. 15, 197-201, 1990

A:Title: Complete sequence of a cDNA of alpha subunit of soybean beta-conglycinin.

A:Reference number: S14681; MUID:91355860

A:Accession: S14681

A:Molecule type: mRNA

A:Residues: 1-605 <SRB>

A:Cross-references: EMBL:X17698; NID:g18535; PIDN:CA35691.1; PID:g18536

R:Shutov, A.D.; Kakhovskaya, I.A.; Besttygina, A.S.; Bulnaga, V.P.; Horstmann, C.; Mu

Eur. J. Biochem. 241, 221-228, 1996

A:Title: Limited proteolysis of beta-conglycinin and glycinin, the 7S and 11S storage

A:Reference number: S74123; MUID:97054613

A:Accession: S74124

A:Molecule type: protein

A:Residues: 189-196, 'H', 198, 'N', 200, 'X', 202-203; 397-408, 'X', 410, 'X', 412-417, 'X', 419-4

A:Experimental source: seed

C:Superfamily: glycinin

C:Keywords: glycoprotein; seed; storage protein

F:1-22/Domain: signal sequence #status predicted <SIG>

F:23-65/Domain: propeptide #status predicted <PRO>

F:63-605/Product: beta-conglycinin alpha chain #status predicted <MAT>

F:261,517/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 315; DB 1; Length 605;

Best local Similarity 100.0%; Pred. No. 1, 1e-26;

Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ENPKHNKCLQSCNSERSYRNOACHARNCLKYKECEGEIPRRPRQHPER 55

Db 31 ENPKHNKCLQSCNSERSYRNOACHARNCLKYKECEGEIPRRPRQHPER 85

RESULT 2

S20007
beta-conglycinin alpha chain precursor - soybean

C:Species: Glycine max (soybean)

C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997

C:Accession: S20007

R:Leleuvre, J.M.; Dickinson, C.D.; Dickinson, L.A.; Nielsen, N.C.

Plant Mol. Biol. 18, 259-274, 1992

A:Title: Synthesis and assembly of soybean beta-conglycinin in vitro.

A:Reference number: S20007; MUID:92119248

A:Accession: S20007

A>Status: preliminary

A:Molecule type: mRNA

C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
C:Keywords: zinc finger
E:55-300/Domain: erba transforming protein homology <ERBA>

```

Qy      11 SCNSERD-----SYRNACNAR---CNLLVKEECCGEEGIIIPRRPRP 50
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      92 TCRNRQDQIDQHHNRNQCQIRLKKCRVGGKRRFAVDNRGRIIPALPEP 139
      Matches 16; Conservative      6; Mismatches      18; Indels

```

RESULT 12
 T00708
 violaxanthin di-epoxidase homolog F22013.3 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 22-Oct-1999
 C:Accession: T00708
 R:Shinn, P.; Buehler, E.; Dewar, K.; Feng, J.; Kim, C.; Li, Y.; Sun, H.; Conway, A.; Col-
 eologis, A.; Ecker, J.R.
 submitted to the EMBL Data Library, April 1998
 A:Description: Genomic sequence for Arabidopsis thaliana BAC F22013.
 A:Reference number: Z14200
 A:Accession: T00708
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-462 <SH1>
 A:Cross-references: EMBL:AC003981; NID:q3063438; PID:q3063441; GSPDB:GNO0059; ATSP:F2201
 A:Experimental source: cultivar Columbia
 C:Genetics:
 A:Gene: ATSP:F22013.3
 A:Map position: 1
 A:Introns: 72/3; 128/2; 160/3; 292/2

Query Match	19.7%	Score 62	DB 2	Length 462																																													
Best Local Similarity	27.6%	Pred. NO. 19																																															
Matches	16	Conservative	7	Mismatches 15; Indels 20; Gaps 3.																																													
QY	8	C	L	O	S	E	N	S	E	R	D	S	T	R	N	O	A	C	A	R	C	-----	N	L	K	A	Y	K	E	K	C	-----	E	E	G	I	P	P	R	P	48								
Db	146	C	L	O	T	T	N	N	P	D	-----	E	T	E	C	O	K	C	O	D	E	F	E	N	S	V	D	E	F	N	C	A	S	R	K	C	V	R	K	S	D	L	G	E	F	P	A	D	200

```

RESULT 13
T26225
hypothetical protein W06D11.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T26225
R:Lloyd, C.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z20176
A:Accession: T26225
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-147 <MIL>
A:Cross-references: EMBL:269718; PIDN:CAA93537.1; GSPDB:GNO0028; CESP:W06D11.5
A:Experimental source: clone W06D11
C:Genetics:
A:Gene: CESP:W06D11.5
A:Map position: X
A:Introns: 103/3

```

```

Query Match      19.4%  Score 61;  DB 2;  Length 147;
Best Local Similarity  40.9%  Pred. No. 8.9;
Matches 18;  Conservative 4;  PmMatches 18;  Indels 4;  Gaps 1
QY  9  LQSCNSERDYSRNQCHARNCLLTKYEKEC-----DEGLIPRRP 48
    | | | | | : | : | | | | | : | : | | | | |

```

Db 50 IMSCLSRRESFVLQEVKDRQMLLKVGLKICNEMLQEGNSNRPTP 93

RESULT 14
T48814
hypothetical protein 1566.220 [imported] - *Neurospora crassa*
C:Species: *Neurospora crassa*
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 05-May-2000
C:Accession: T48814
R:Schulze, U.; Alyn, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakaturu
submitted to the Protein Sequence Database, April 2000
#:Reference number: 224541

A:Residues:1-1952 <SCH>
A:Cross-references: EMBL:AL358822; GSPDB:GM00112; NCSP:15E6.220
A:Experimental source: cosmid contig 15E6, strain 74
C:Genetics:
A:Gene: NCSP:15E6.220
A:Map position: 2
A:Introns: 281/3

Query Match	19.4%	Score 61	DB 2	Length 1952
Best Local Similarity	24.7%	Pred. No. 85		
Matches 18	Conservative 16	Mismatches 19	Indels 20	Gaps 3

```
OY      2 NPKHNKCSLOSNSRDSYSRNQACH-----ARCLIK-VEKEECC-----G 41
          :|: |: |::|||           ||:::||||| ||
Db     585 SPEDNEVMASVDPRRDSSQSLSERPEDILLPRAPRYHMETYEKERAEDEDDARSQS 644
OY      42 EIPRRPPRPQHHE 54
          :|: |||: |
Db     645 QSVKPEEPREPEPD 657
```

RESULT 15

JH0422

voltage-dependent calcium channel complex alpha-1 chain - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 11-Jan-2000

C:Accession: JH0422; D35901; I60901

R:Hui, A.; Ellnor, P.T.; Krizanova, O.; Wang, J.J.; Diebold, R.J.; Schwartz, A.

Neuron 7, 35-44, 1991

A:Title: Molecular cloning of multiple subtypes of a novel rat brain isoform of the a

:Reference number: JH0422; MUID:91299338

A:Molecule type: mRNA
A:Residues: 1-1646 <NU>
A:Cross-references: GB:K57682; NID:g206573; PIDN:ANAA2015.1; PID:g206574
A:Experimental source: Brain
R:Snutch, T.P.; Leonard, J.P.; Gilbert, M.M.; Lester, H.A.; Davidson, N.
Proc. Natl. Acad. Sci. U.S.A. 87, 3391-3395, 1990
A:Title: Rat brain expresses a heterogeneous family of calcium channels.
A:Reference number: A35901; MUID:90239020
A:Accession: D35901
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A:Molecule type: mRNA
A:Residues: 1247-1434 <SNU>
A:Experimental source: brain
R:Yu, A.S.L.; Hebert, S.C.; Brenner, B.M.; Lytton, J.
Proc. Natl. Acad. Sci. U.S.A. 89, 10494-10498, 1992
A:Title: Molecular characterization and nephron distribution of a family of transcrip
A:Reference number: A46422; MUID:93066265

A:Status: preliminary: translated from CB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1040-1261,1305-1365 <RES>
A:Cross-references: GB:K99221; NID:g203370; PIDN:AAA0895.1; PID:g203371
A:Experimental source: kidney
A:Comment: Calcium channels are essential for many cellular functions, such as muscle

C:Superfamily: voltage-dependent calcium channel protein alpha-1 chain
C:Keywords: alternative splicing; calcium binding; calcium channel; glycoprotein; phosph
F:1463-1491/Domain: calcium binding #status predicted <EFC>
F:154,224,328/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:464,848,1489,1584/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 19.0%; Score 60; DB 2; Length 1646;
Best local similarity 34.9%; Pred. No. 94;
Matches 22; Conservative 9; Mismatches 18; Indels 14; Gaps 4;
QY 1 ENPKHNK--CLQCSNTER---DSTRNQA---CHARCNLLKVEKECEGEGLIPR----P 46
DB 783 ENKKNNKPEVNOIANSNDKVTIDYQEEAEEDKDPYPCDVPVGEVEEEEEEDEDEPEVPAGP 842
QY 47 RPR 49
DB 843 RPR 845

Search completed: March 1, 2001, 16:10:46
Job time: 1077 sec

